In many cases, sigmoid function or logistic distribution is widely used in many research when we try to fit model for category dependent variables. But why we use them often? Because there are many alteration of it like hyper tangent function or else. This can be proved by Chernoff bound.

Proposition1)Let X be arbitrary random variable(r.v) and $a, t > 0$ be real number.

$$P(X > a) \leq \frac{E(e^{tX})}{e^{ta}}$$

Proof) If we set $Y = e^{tX}$ and use Markov's inequality, we can get it.

Suppose $a' = \frac{\ln(1+e^{ta})}{t} > 0$ and $E(e^{tX}) = 1$. If we change a as a', then $P(X > a) \leq \frac{E(e^{tX})}{1+e^{ta}}$ holds since proposition1 is true for any a. If we assume $\left(1 - E(e^{tX})\right) > 0$, then

$$P(X \leq a) = 1 - P(X > a) \geq 1 - \frac{E(e^{tX})}{1+e^{ta}} = \frac{e^{ta} + \left(1 - E(e^{tX})\right)}{1+e^{ta}}$$

$$\geq \frac{e^{ta}}{1+e^{ta}} \dots (1)$$

The (1) is called logistic distribution and since X, t are arbitrary so we can use X, t as $\log\left(ratio\left(\frac{(X=1|V)}{(X=0|V)}\right)\right), V\beta$ for each. If $E(e^{tX}) = e^{V\beta}E(e^{X}) = e^{V\beta}E\left(ratio\left(\frac{(X=1|V)}{(X=0|V)}\right)\right) < 1$, $E\left(ratio\left(\frac{(X=1|V)}{(X=0|V)}\right)\right) < e^{-V\beta}$, then we can use logistic distribution for probability. I'll give an example for this explanation.

Suppose $X = 1$ is one people's survival state and $X = 0$ is death state, V is a amount of one's smoke in a week and β is a effect of V. Then we can think easily that as $ratio\left(\frac{(X=1|V)}{(X=0|V)}\right)$ is smaller as V get larger, but if we find that ratio is exponential then we can use logistic distribution for probability. This means that if we don't know about distribution of some r.v but only exponential convergence rate of ratio for binary variable, we can use logistic distribution in worst case. But if we have domain knowledge then we should use it since the logistic regression may be inefficient in this case.